Realizing the Virtues (and Combating the Vices) of the Digital Shift in Heritage Scholarship: the Role of Knowledge Representation

George Bruseker (Takin.solutions, CIDOC CRM SIG)

FAIR Heritage: Digital Methods, Scholarly Editing and Tools for Cultural and Natural Heritage (Tours / Athens / World / Zoom) 17/6/2020

Plan of Action

- 1) FAIR Principles and Knowledge Representation (KR)
- Achievements in KR to support the goals of Interoperability and Reusability (IR)
- 3) Challenges in pushing forward KR for IR

Fair Principles and Knowledge Representation

What is FAIR?

- A set of principles
- Recognition of a problem (beyond the epistemic / technic)
- A call for justice



Image: https://book.fosteropenscience.eu/

FAIR appeals to a universalist vision of scholarship and science, a humanistic, communally organized search for the truth that builds itself on the free exchange of information, towards the collective advancement of human knowledge

What is unFAIR?

- A breach of principles
- Ignorance of a problem (the ethical aspect not seen)
- A lack of justice



We experience the present situation as unFAIR because there is a commonly shared view of the basis for a just exchange of information which is implicitly appealed to and expected by scholars and scientists and which the digital transformation and information society promises explicitly and implicitly to bring about, but that it cannot or will not do

The Road to unFAIR: The Information Superhighway



"As we enter this new millennium, we are learning a new language. It will be the lingua franca of the new age. It is made up of ones and zeros and bits and bytes. But as we master it ... as we bring the digital revolution into our homes and schools ... we will be able to communicate ideas, and information -- in fact, entire Toni Morrison novels -- with an ease never before thought possible.

...after the next Big Bang, in the ensuing expansion of the information business, the new marketplace will no longer be divided along current sectoral lines. There may not be cable companies or phone companies or computer companies, as such. Everyone will be in the bit business. The functions provided will define the marketplace. There will be information conduits, information providers, information appliances and information consumers."

Remarks by Vice President Al Gore (as prepared) Vice-President Al Gore Royce Hall, UCLA, Los Angeles, California January 11, 1994 <u>https://web.archive.org/web/20060929192826/http://www.clintonfoundation.org/leg</u> acy/011194-remarks-by-the-vp-on-television.htm

Communication

Openness/Freedom



Equity

ICT 1990s to Now: Individuals Using the Internet

Individuals using the Internet (% of population)

International Telecommunication Union, World Telecommunication/ICT Development Report and database.



1994: 0.446% of the Population

2017: 49.723% of the Population

Source: World Bank

https://data.worldbank.org/indicator/IT.NET.USER.ZS?end=2018&

start=1994&view=chart

ICT 1990s to Now: Mobile Cellular Subscriptions

Mobile cellular subscriptions

International Telecommunication Union, World Telecommunication/ICT Development Report and database.



1994: 55 Million

2017: 7.75 Billion

Source: World Bank

https://data.worldbank.org/indicator/IT.CEL.SETS

ICT 1990s to Now: Transistors on an Integrated Chip



Data source: Wikipedia (https://en.wikipedia.org/wiki/Transistor_count)

The data visualization is available at OurWorldinData.org. There you find more visualizations and research on this topic.

Licensed under CC-BY-SA by the author Max Roser



What does the call to FAIRness mean?

- European Open Science Declaration -> the revolution called for again... can we do it right?
- Critically address several decades down the path of a transformation towards a digital, information society
- A moment to critically assess the benefits and pitfalls of the tools we have built -> evaluated by the criterion of the principle of justice we have in mind
- Consider how to develop new ideas of how to realize justice that are adequate to a digital / information society
- A chance for the humanities in particular to offer their tools of critical assessment in order to move the agenda forward

FAIR Focus: Interoperability and Reusability

What enables interoperability and reusability? -> Comprehensibility.

Comprehensibility is not a given but gained, developed through language and through developing disciplinary codes and means of relaying information.

Many of the vices that we see in the implementation of the information society relate to a lack of focus on supporting human comprehension in the digital space





Knowledge Representation a central key to FAIRness



In an information society, your data is one of the fundamental ways of expressing yourself both as an individual and scientifically and yet:

- we lack the formalized languages and codes to make this expression comprehensible and interrogable
- we lack the automated tools to connect to these expressions and explore them in a way which makes them tractable to comprehension and the expansion of knowledge

Building digital lingua franca(s) for a diverse audience



Domain Specialist Software Developer Conceptual Modeller Machines

KR as the Answer

- "An ontology is a formal, explicit specification of a shared conceptualization."
 Studer, 1998
- Machine readable, human understandable
- A collaborative effort to:
 - Create formal models (ontologies) for expressing analytic knowledge consistently
 - Adopt extant formal models (ontologies) in order to express and share analytic knowledge consistently



Ethics and Argument in Plato's Republic. Summer, 1994, Steven S. Tigner & Tonya Geckle.

Achievements in KR to support the goals of Interoperability and Reusability (IR)

Plans for a Semantic Web of Museum Knowledge

conseil ICON conseil international des musées



"At the ICOM meeting in Stavanger, the Data Model Working Group and the Data Terminology Working Group merged and became the Documentation Standards Working Group...

Future CIDOC data models will be developed using O-O methods, techniques, and formats.

The O-O model will retain all the related information contained in the existing relational model, and it will provide flexibility and extensibility not possible for a relational approach.

The O-O model will enable more specific collections management and research information to be included. **Communication, interchange of information, and** public access to museum information will be enhanced.

The O-O model requires a different mind-set for looking at museum data: although it is more flexible and extensible, it also is more complex."

Documentation Standards Working Group Interim Report, ICOM, Stavanger, April 1996

Achievements in KR for FAIR Data



Knowledge Representation

BUILDING ONTOLOGIES WITH BASIC FORMAL ONTOLOGY

ROBERT ARP | BARRY SMITH | ANDREW D. SPEAR







Ontological Models







Vocabularies and Reference Data



Forum on Information Standards in Heritage





















Linked Art

Mémoires des Archéologues













et des Sites Archéologiques

CONCEPTUAL



Data for History

Modelling, curation, interoperability

ARIADNE







Tools



Extract, Transform, Load (ETL) and Explore

Process



Challenges in pushing forward KR for IR

Where we need to go



Building Understanding

Problem

You can't have a lingua franca without (competent) users

Solution

You need to create curricula and material that addresses the learning needs and approaches of various user groups and allows them to actually communicate







Semantic Reference Data Models



Element Field Collection			
Identifier Attribution	Identifier	This field is used to indicate an identifier attributed to the documented person.	\rightarrow P1 \rightarrow E42[1]
Identifier Attribution	ldentifier Type	This field is used to indicate the type of an identifier attributed to the documented person.	\rightarrow P1 \rightarrow E42[1] \rightarrow P2 \rightarrow E55["Identifier Type"]
ldentifier Attribution	ldentifier Provider	This field is used to indicate the provider of the identifier attributed to the documented person	\rightarrow P1 \rightarrow E42[1] \rightarrow P37i \rightarrow E15 \rightarrow P14 \rightarrow E39
Identifier Attribution	Identifier Source	This field is used to indicate the source based on which the identifer was attributed to the object.	\rightarrow P1 \rightarrow E42[1] \rightarrow P37i \rightarrow E15 \rightarrow P16 \rightarrow E73
n/a	Name	This field is used to indicate the main name attributed to the documented person. No part break down of name is here implied. The full name of an individual is expected here where available.	$\rightarrow P1 \rightarrow E41[2] \bullet$ $\rightarrow P1 \rightarrow E41[2]$ $\rightarrow P2 \rightarrow$ E55['Preferred Name']
n/a	Name Part	This field is used to indicate a type of a specific part of the name attributed to the documented entity.	$\rightarrow P1 \rightarrow E41[2]$ $\rightarrow P106 \rightarrow$ $E41[22] \rightarrow P2 \rightarrow$ E55
n/a	Name Language	This field is used to indicate the language of the name attributed to the documented entity.	\rightarrow P1 \rightarrow E41[2] \rightarrow P72 \rightarrow E56

https://docs.swissartresearch.net/

Teaching Ontologies as a Language





Maintain and Develop Communities

Problem

For semantics to work they need to be adopted by their audience in their practice



Solution

Establish long term institutional and academic support

Recognize work and output

Engage the community

Open the community

Invite Diversity

Be transparent with the community



Linked Art Community

Linked.Art is a community project in which all interested parties are welcome to participate, according to the <u>code of conduct</u>.

Get Involved

There are many ways to get involved:

- Attend a Linked Art Event
- Join the Discussion Group: <u>https://groups.google.com/forum/#!forum/linked-art</u>
- Join the Slack (email <u>rsanderson@getty.edu</u> for an invitation)
- Add to Issues / Questions: https://github.com/linked-art/linked.art/issues
- Edit the website source in Github: <u>https://github.com/linked-art/linked.art/</u>
- Listen to our recorded discussions: <u>https://www.youtube.com/channel/UCNASnutgByTdQHGehoOUISA</u>



Expand Expressivity

Problem

Semantics should enable you to say more and do more with your data, but applications are built to mimic existing functionality

Solution

Break out of existing paradigms of databases, explore new ways of organizing and visualizing information (while maintaining connection to known, authoritative data structures)





https://www.researchspace.org/



ETAT 3, DATATION ET FONCTION

P0_12. Dans l'Etat 3, le mur pignon ouest (M42) et le mur nord (M35) sont détruits : le bâtiment 11 devient plus long et plus étroit. SECTION 4 - LE PRESBYTÈRE DE RIGNY ET SES DÉPENDANCES (MILIEU 15E-MILIEU 19E

[-] Le bâtiment 11

s.)

- [+] Etat 1, datation et fonction
- [+] Etat 2, datation et fonction
- [+] Etat 3. datation et fonction
- [-] Le bâtiment 9

- [+] Transformation du centre paroissial
- [+] Le bâtiment 5
- [+] Le bâtiment 8

Circuler dans les illustrations

Accéder au diagramme logiciste



Commentaire

Le nouveau mur nord (M36) du bâtiment 11 est construit en retrait de l'ancien, dans le prolongement du mur nord (M85) du bâtiment 22. Le four construit à l'ouest, sur l'emprise du bâtiment 22, prend appui sur le conduit des anciennes latrines du bâtiment 11 et scelle la tranchée de destruction de l'ancien mur pignon M42.

P0_13. D'après un procès-verbal de 1824, le bâtiment 11 sert de cuisine au nouveau presbytère construit plus au nord en 1822 (bâtiment 5).

Commentaire

« Le conseil... a reconnu que la réclamation de M. le desservant était juste ... et a examiné que le four aui est en dehors est en très mauvais état sans pouvoir cuire le pain attendu au'il recoit toute l'eau de la toiture... Le conseil a aussi remaraué aue le plancher de ladite cuisine est très vieux, au'il fond de toutes parts et qu'il est urgent de la refaire à neuf » (A.D.I.L., série D, conseil municipal du 8 mai 1824).

P1 3. L'étage du bâtiment 11 est supprimé.

https://www.unicaen.fr/puc/rigny//accueil

https://www.researchgate.net/publication/331 293193 The Archaeological Excavation Re port of Rigny An Example of an Interoper able Logicist Publication



Olivier Marlet, Elisabeth Zadora-Rio, Pierre-Yves Buard, Béatrice Markhof, and Xavier Rodier

[+] Etat 1. datation et fonction

[+] Etat 2. datation et fonction



Continuous, Responsible Revision of the Foundation

Problem

Ontology design is open ended on purpose, allowing for the expansion and revision of knowledge



Solution

Openness to Revision

Build bridges between communities (Diversity)

Cultivate spirit of collaboration and not competition

OntoME

Ontology Management Environment - beta version							Data fo	ata for History Consortium				
Home Classes -	Properties	Namespaces	Projects	Profiles	O User guide							
Classes											Complex	
Show 50 v entries											Search:	
Class name				↓≞ Nar	nespace					11	Last updated	J↑
Acquisition – E8				CID	OC CRM						2018-11-16	
Activity – E7				CID	OC CRM						2018-11-16	
Actor Appellation - E82				CID	OC CRM						2018-11-16	
Actor – E39				CID	OC CRM						2018-11-16	
Address – E45				CID	OC CRM						2018-11-16	
Aggregation Work - F17				FRE	Roo: Functional	Requirements for Bibliographic	Records				2018-11-16	
Alteration - S18				CRM	Asci: An Extensio	on of CIDOC CRM to support sc	cientific observation				2019-01-12	
Amount of Fluid – S12				CRM	Asci: An Extensio	on of CIDOC CRM to support sc	cientific observation				2019-01-12	
Amount of Matter - S11				CRM	Asci: An Extensio	on of CIDOC CRM to support so	cientific observation				2019-01-12	
Annotation Event - D30				CRM	/Idig: An Extensio	on of CIDOC CRM to support pr	rovenance metadata				2019-01-25	
Annotation Object - D29				CRM	/Idig: An Extensio	on of CIDOC CRM to support pr	rovenance metadata				2019-01-25	
Appellation – E41				CID	OC CRM						2018-11-16	
Archaeological Excavation	ı – A9			CRM	/larchaeo: An Ext	ension of CIDOC CRM to supp	ort the archaeological exc	cavation process			2018-11-16	
Area – D35				CRM	/Idig: An Extensio	on of CIDOC CRM to support pr	rovenance metadata				2019-01-25	
Argumentation - I1				CR	/linf: An Extensio	n of CIDOC-CRM to support ar	gumentation				2019-12-04	

http://ontome.dataforhistory.org/

Browse

Browse the library of ontologies (2) Search... Showing 866 of 1040 Sort: Popular Submit New **Current Procedural Terminology (CPT)** Ontology Current Procedural Terminology Uploaded: 11/18/19 Entry Type Ontology (866) Medical Dictionary for Regulatory Activities Terminology Ontology View (174) (MedDRA) (MEDDRA) MedDRA is an international medical terminology with an emphasis on use for data entry, Uploaded in the retrieval, analysis, and display Last Uploaded: 11/18/19 ~ SNOMED CT (SNOMEDCT) Category **SNOMED Clinical Terms** All Organisms (31) Uploaded: 11/18/19 Anatomy (72) □ Animal Development (15) **RxNORM (RXNORM)** □ Animal Gross Anatomy ... **RxNorm Vocabulary** Arabidopsis (3) Uploaded: 11/18/19 Biological Process (50) ___. . . <u>.</u> **Online Mendelian Inheritance in Man (OMIM)** Group Online Mendelian Inheritance in Man, OMIM (TM) BIBLIO (10) Uploaded: 11/18/19 BIS (3)

http://bioportal.bioontology.org/

KR's (Potential) Contribution to Establishing FAIRness

Can help support the virtues and not the vices of the digital transform by building new, common digital languages and new forms of digital literacy

Communication: common, community owned languages to speak **Equity**: community grounded languages to speak in

Openness/Freedom: ground up open source mentality, open is the driving motivator

Knowledge: making information tractable to knowledge communities, empowering them to understand and query information at a new scale

Simplicity: ... exploiting the power of reasoning and patterns to re-ignite the idea of smart agents

Thank you Your Feedback Welcome

George Bruseker george@takin.solutions